© August 2023| IJIRT | Volume 10 Issue 3 | ISSN: 2349-6002

Signal Processing Techniques for Drone Detection

Manan Jandial[1], Teekshan Heera[1], Deepak Gupta[1], Bhawna Sharma[2], Sheetal Gandotra[3]
[1]*Student, Government College of Engineering and Technology Jammu, J&K, India.*
[2]*Professor, Government College of Engineering and Technology Jammu, J&K, India.*
[3]*Assoc Professor, Government College of Engineering and Technology Jammu, J&K, India*

***Abstract*— Drones have garnered great popularity due to their widespread applications across various industries, ranging from surveillance and reconnaissance to logistics and entertainment. However, the increasing use of drones has also given platform to potential security threat activities, such as unauthorized surveillance, smuggling, and privacy invasion. As a result, there is need to develop efficient and reliable drone detection systems to safeguard sensitive areas and public spaces. This paper focuses on the development and implementation of an inclusive drone detection system based on advanced signal processing techniques and machine learning algorithms and deep learning algorithms. The proposed system aims to identify and classify drones with highest accuracy on real world datasets. Using the real-world datasets and different highly accurate techniques for classification the model developed is reliable and can be directly used for advance practical purposes.**

I. INTRODUCTION

The rapid proliferation of drones has revolutionized numerous industries and introduced a myriad of innovative applications. From aerial photography and environmental monitoring to package delivery and disaster response, drones have become an integral part of modern-day society. However, widespread adoption of drone also raises concerns about potential security threats and safety risks [1]. Instances of unauthorized surveillance, privacy invasion, and the misuse of drones in restricted areas have prompted the urgent need for effective drone detection systems. This project endeavors to address the challenges concerned with drone detection by leveraging the power of advanced signal processing techniques and machine learning algorithms. By combining the strengths of both disciplines, the aim is to develop a comprehensive and accurate drone detection system capable of distinguishing between drone signals and other sources of noise in various environments. The introduction of drones has significantly diversified the sources of signals in the electromagnetic spectrum. Unlike traditional radar systems that deal primarily with aircraft and terrestrial objects, drone detection requires a more nuanced approach due to the distinctive characteristics of drone signals. Three common ways to for drone detection include RF detection, visual detection and acoustic detection. Drones emit signals across different frequencies, including radio frequency (RF) signals used for communication and control, acoustic emissions generated by their propellers and engine, and even visual cues from their flight in the sky. This [2] research provide us with great in-depth knowledge about RF detection. Shi [3] introduced an innovative approach involving Hash Fingerprint features and distance-based support vector data description (SVDD) techniques. This methodology was devised for the identification of slow-moving, compact unmanned aerial vehicles (LSSUAVs) operating within the 2.4 GHz frequency range. This research [4] gives us insight into the visual detection of drones. Saqib [5] investigated different pre-trained models including VGG16 with the Faster R-CNN model for the detection of drones from video data. Our research will commence by focusing on the acoustic part of the drone. collecting and preprocessing diverse datasets containing drone and non-drone audio signals. This preprocessed data will be crucial for training and evaluating the machine learning model, ensuring its ability to differentiate between legitimate drone activity and false positives from environmental noise or other benign sources. In the subsequent stages, advanced acoustics processing techniques will be applied to extract pertinent features from the collected signals. The usage of time-frequency analysis, wavelet transforms, and spectrogram-based methods will aid in capturing the unique modulation patterns and acoustic signatures characteristic of drone signals. These

features will serve as vital discriminative factors for the subsequent machine learning model. Machine learning algorithms, 2D and 1D Convolutional Neural Networks (CNNs) and Artificial Neural Network (ANNs), will be used to learn and understand the complex patterns inherent in drone signals. The deep learning model will be fine-tuned and optimized using appropriate loss functions to achieve highest accuracy possible in distinguishing between Drone and non-drones.
The goal of this project is to develop a robust, efficient, and real-time drone detection system that can be deployed in various settings, including urban areas, critical infrastructure facilities, and sensitive zones. Such a detection system will significantly bolster security measures, safeguard privacy, and protect against potential drone-related threats. By combining advanced signal processing with deep learning, this research endeavor seeks to contribute to the growing body of knowledge in drone detection technology. As drones continue to evolve and find new applications, this project aims to lay the base for further advancements and scope in the field of drone detection, ensuring a safer and more secure and peaceful future for our technologically interconnected world.

1.1 Objectives of the Study:

The main objective of this study is to develop a vigorous and efficient drone detection system that utilizes advanced signal processing techniques and machine learning algorithms. The specific scope of the study include:

Drone Signal Classification: To accurately classify drone signals and sound, from various sources, such as RF communication, acoustic emissions, and visual cues, and differentiate them from other noise sources present in the environment.

Signal Preprocessing: To collect and preprocess diverse datasets containing drone and non-drone signals, ensuring that the main data used for training and testing the model is representative of real-world scenarios.

Feature Extraction: To employ advanced signal processing methods, like time-frequency analysis, wavelet transforms, and spectrogram-based techniques, to extract relevant features from the drone signals that are distinctive and informative for discrimination.

Deep Learning Model: To create a deep learning model, utilizing Convolutional Neural Networks (CNNs) and Artificial Neural Networks (ANNs), capable of learning complex patterns present in drone signals and achieving high accuracy in detection.

Model Optimization: To fine-tune and optimize, improve the machine learning model using appropriate loss functions, regularization techniques, and hyperparameter tuning to enhance and improve its performance and generalization capabilities.

Real-World Deployment: To test and evaluate the developed drone detection system in diverse environments, such as urban areas, critical infrastructure facilities, and sensitive zones, to assess and evaluate its performance under different conditions.

Security Enhancement: To contribute to enhancing security measures by providing a reliable drone detection system capable of safeguarding sensitive areas and mitigating potential security threats posed by drones.

Privacy Protection: To aid in protecting privacy by detecting and preventing unauthorized drone activities that may impose upon individuals' privacy rights.

Advancement in Technology: To give and contribute to the advancement of signal processing and machine learning techniques in the domain of drone detection, paving the way for future scope and research and development in this rapidly evolving field.

By achieving these objectives, the study aims to provide a comprehensive solution to the pressing need for optimized and accurate drone detection systems, ultimately ensuring a safer and more secure environment in the face of increasing drone usage.

II. DATA ACQUISITION

In the data acquisition phase of this study, we sourced the necessary datasets to facilitate our research objectives. To harness spatial information, we obtained drone-captured datasets from a publicly available repository on GitHub [6]. These drone datasets, generated by previous researchers, encompass diverse geographical locations, and exhibit varying flight parameters, ensuring the breadth of data required for our analysis.

To assess the impact of noise on our study, we procured noise dataset from Pixabay, a reputable online platform for high-quality audio content. The noise captured various scenarios, allowing us to simulate different noise levels within our study environment. By incorporating these external data we aimed to create a comprehensive and realistic representation of noise effects in our analysis.

III. DATA PREPROCESSING TECHNIQUES

Digital Signal Processing Concepts in Aerial Acoustics Detection

In the quest to delve into the world of Aerial Acoustics Detection using Machine Learning, a solid foundation in Digital Signal Processing (DSP) concepts becomes imperative. This subsection aims to provide an overview of the main components and features of sound that lay the groundwork for the efficient and successful implementation of the drone detection project. The Audios wave plots to understand the sound better with the comparison of all audios are plotted below:

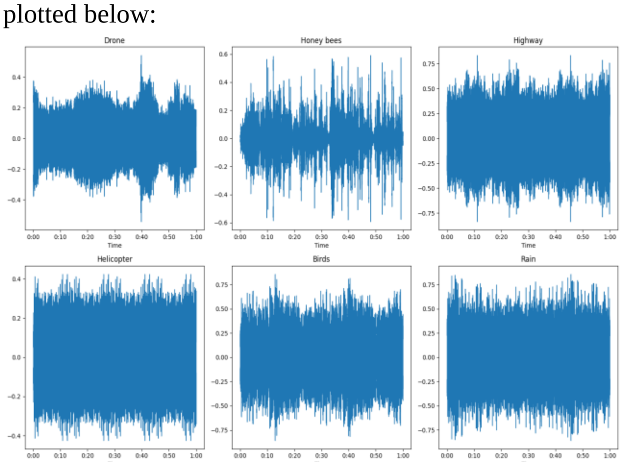


Figure 1: Comparative wave plots of various audio

1. Understanding Sound Basics: Amplitude, Pitch, Frequency, and Waveform.
At the core of digital signal processing lies a profound comprehension of fundamental sound properties. Amplitude represents the magnitude or intensity of a sound wave, determining its loudness. Pitch denotes the perceived highness or lowness of a sound and is determined by the frequency of the sound wave. Frequency refers and point to the number of oscillations per unit of time and governs the pitch and intensity of a sound. Lastly, waveform illustrates the shape of a sound wave and plays an important role in identifying unique acoustic signatures emitted by drones.

2. Time Domain Features: Unraveling Sound Characteristics in Time.
Time domain features encapsulate vital attributes of sound signals over time. Among the significant features used in Aerial Acoustics Detection are:

a. Amplitude Envelope: The amplitude envelope outlines the varying amplitude of a sound wave over time, revealing crucial temporal information about the sound's structure.

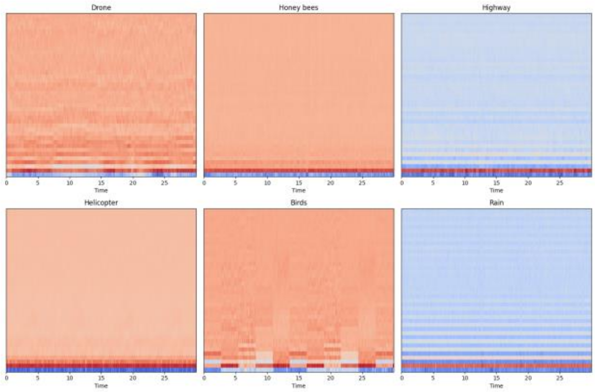
b. Root Mean Square (RMS) Energy: RMS Energy calculates the average energy of a sound signal within a particular time frame, allowing us to understand the overall power distribution.

c. Zero Crossing Rate: This feature characterizes the rate at which the signal changes polarity (from positive to negative and vice versa) and is particularly useful in differentiating between periodic and noisy signals.

3. Time-Frequency Domain Features: Capturing Time-Varying Spectral Information.
Time-frequency domain features are essential in grasping the dynamic nature of sound signals. In the context of drone detection, the following features are particularly relevant:

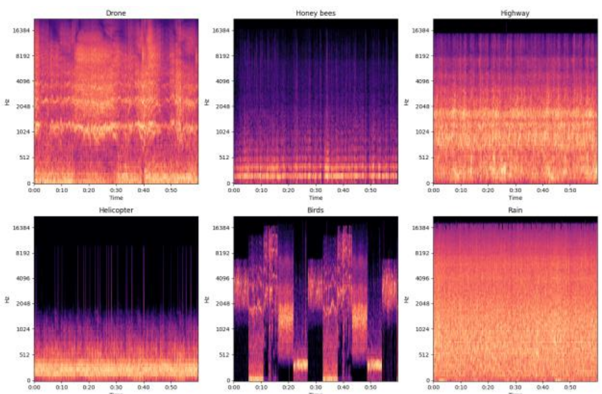
a. Mel Frequency Cepstral Coefficients (MFCC): MFCC represents the short-term power spectrum of a sound, effectively capturing its spectral characteristics. These coefficients and features are widely used in speech and audio processing tasks, including sound classification.

The plot for the MFCC of all the audios that we have used is plotted below for better representation. The MFCC calculated by using frame length equals to 1024 and hop length equals to 512 with number of mels equals to 20.



B. Mel Spectrogram: A visual representation of the spectrum of frequencies of a sound signal as it varies with time. Spectrograms help in identifying time-varying frequency components, distinguishing drones from other environmental sounds.

The mel-spectrogram plot for the audios we are using for better representation:



c. Frequency Domain Features: Further analyzing the frequency content of the sound, features like Spectral Centroid and Spectral Bandwidth provide insights into the spectral distribution and bandwidth of the signal.

d. Band Energy Ratio: This feature quantifies the energy distribution across different and various frequency bands, allowing for efficient discrimination of drone sounds from background noises.

IV. DATA FEATURE EXTRACTION

In the dynamic landscape of Aerial Acoustics Detection using Machine Learning, the effective preprocessing of audio data is crucial for extracting meaningful insights and accurate classification of drone sounds. Python, a versatile and widely-used programming language, serves as the support for the entire data processing pipeline. In conjunction with the powerful "Librosa" library, which boasts a plethora of inbuilt functions for feature extraction, we harnessed the full potential of these tools to conduct research with utmost precision and efficiency.

1. Python: The Versatile Language Powering the Data Pipeline Python, with its simplicity and rich ecosystem of libraries, has become a preferred choice for data scientists and researchers. Its ability to seamlessly integrate with other technologies makes it an ideal language for handling audio data. Throughout the project, we utilized Python to perform diverse tasks, such as data ingestion, manipulation, visualization, and model development. The language's adaptability enabled us to prototype and implement various preprocessing techniques with ease, facilitating the progression of the research project.

2. Librosa Library: Empowering Feature Extraction with Ease at the heart of the data preprocessing phase lies the "librosa" library, a powerful Python package explicitly designed for music and audio analysis. With its vast array of built-in functions, the library simplifies the extraction of key features from audio signals. Leveraging librosa's capabilities, we significantly reduced the complexity of the data preprocessing process.

3. Feature Extraction and Analysis: The "librosa" library provided a range of feature extraction functions, allowing us to capture the unique characteristics of drone sounds from raw audio data. These features included the time domain features, such as amplitude envelope, root mean square energy, and zero crossing rate, which provide important insights and information into the temporal aspects of the sound. Additionally, time-frequency domain features like Mel Frequency Cepstral Coefficients (MFCC) played an important role in capturing the spectral characteristics of drone sounds. By transforming the audio signals into Mel-frequency cepstral domain, we obtained a compact and informative representation that facilitated the subsequent classification process.

5. Implementation and Validation of Test Audios: Upon completing the data preprocessing phase, we

applied the developed feature extraction techniques to a diverse set of test audio samples, specifically focusing on drone sounds. Through comprehensive analysis and validation, we ascertained the effectiveness of the preprocessing tools and confirmed the ability and potential of the feature extraction methods to discern drone sounds from other background noises accurately.

In conclusion, the integration of Python as the programming language and the "librosa" library for feature extraction endowed us with a robust and efficient data preprocessing pipeline. By successfully navigating this critical phase, the groundwork and foundation was laid for further analysis and machine learning-based classification of drone sounds.

V. VARIOUS MODELS USED

This section focusing on machine learning models for drone sound classification. By leveraging the wealth of algorithms available in machine learning, we explored and evaluated different classifiers, seeking the optimal approach for Aerial Acoustics Detection. We delved into popular machine learning algorithms like Support Vector Machines (SVM), Random Forests, k-Nearest Neighbors (k-NN), and Decision Trees. Each algorithm was meticulously tested and tuned to ensure robustness and accuracy in classifying drone sounds against various background noises.

Machine Learning Approach: Support Vector Machine (SVM): The machine learning approach involves leveraging the widely-used Support Vector Machine (SVM) algorithm for binary classification. SVM has proven for having powerful and versatile algorithm in various domains, including drone detection. It excels in scenarios where the data is not linearly separable, making it well-suited for handling complex acoustic signatures of drones. Main primary advantages of SVM is its ability to check and identify optimal hyperplanes that effectively separate drone and non-drone data points in the feature space. This enables SVM to achieve highest accuracy and generalization, even in the presence and enlightenment of noise and overlapping acoustic patterns. Moreover, SVM's binary classification nature simplifies the task of distinguishing between drones and non-drones, making it computationally efficient and straightforward to implement. With proper feature engineering and tuning of hyperparameters, SVM can provide reliable and robust drone classification results.

Convolutional neural networks: Convolutional Neural Networks (CNNs) have revolutionized the field of computer vision and image processing. These deep learning architecture and algorithms have demonstrated remarkable capabilities in tasks such as image classification, object detection, facial recognition, and more. CNNs are designed to mimic the visual processing abilities and performance of the human brain, enabling them to extract meaningful patterns and features from raw input data. We will delve into the fundamental 2-dimensional layers of a CNN and elucidate their functionalities in image processing tasks.

1. Convolutional Layer: The convolutional layer is the cornerstone of CNNs. It performs the essential operation of convolution, where a set of learnable filters (also called kernels) slide over the given image to detect patterns. Each filter extracts specific features, such as edges, textures, or shapes, by computing element-wise multiplications and summations. The output of the convolutional neural network layer is called feature maps, which preserve spatial information and highlight the detected patterns in the input image.

2. Activation Layer: The activation layer is typically applied immediately after the first convolutional layer. It introduces non-linearity into the network, allowing it to learn complex patterns and relationships within the data. The Rectified Linear Unit (ReLU) is commonly used activation function in CNNs. It sets all negative pixel values to zero while retaining positive values, thereby enhancing the model's ability to capture intricate features.

3. Pooling Layer: The pooling layer plays an important role in down-sampling the feature maps. It reduces the spatial dimensions of the data while keeping its essential information. Max-pooling is prevalent pooling technique in CNNs. It partitions the feature maps into small regions (e.g., 2x2 or 3x3) and keeps only the maximum value within each region, discarding the rest. By doing so, pooling reduces computational complexity, increases the model's efficiency, and promotes translational invariance.

4. Batch Normalization Layer: The Batch Normalization layer is employed to address the internal covariate shift problem, wherein the distribution of feature values changes during training. It normalizes the output of every layer by standardizing it to have zero mean and unit variance. This helps stabilize the training process, accelerates convergence, and allows for higher learning rates, leading to faster training and improved model performance.

5. Dropout Layer: The Dropout layer is a regularization technique used to prevent overfitting in CNNs. During training, it randomly drops (sets to zero) a fraction of neurons in the network. This forces the model to learn redundant representations and reduces the dependence on specific neurons, making the network more robust to unseen data.

6. Fully Connected Layer: The fully connected layer is last layer of the CNN, responsible for making predictions based on the learned features. It flattens the output from the previous layers into a one-dimensional vector and passes it through a traditional artificial neural networks. The neurons in this layer are fully interconnected, and their weights and length are updated during training to optimize the model for the specific task at hand, such as image classification or detection.
Convolutional Neural Networks have proven to be a transformative technology in the domain of computer vision. Their ability to learn the hierarchical representations directly from raw pixel data enables them to excel in a wide range of image processing tasks. By understanding the key 2-dimensional layers and their functionalities within CNNs, researchers and practitioners can design more sophisticated architectures and continue to push the boundaries of what's possible in the realm of computer vision and deep learning.

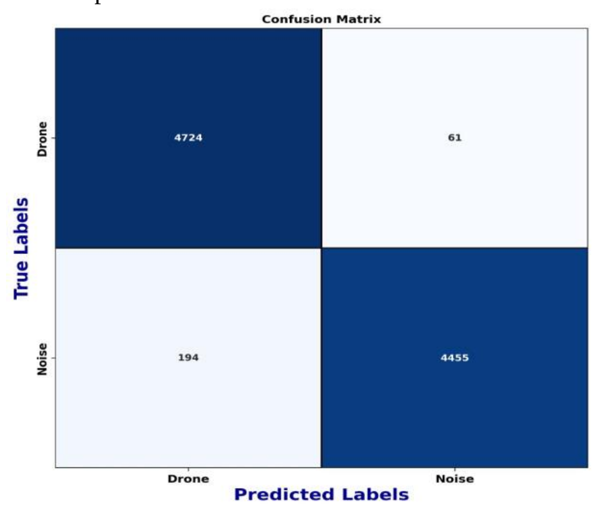
CNN 1D: Convolutional Neural Network (CNN) 1D is a deep learning architecture designed for processing one-dimensional sequential data, like time series or audio signals. Unlike traditional CNNs that work with images, CNN 1D applies 1D convolutional filters to learn relevant patterns and features from the data given as input. The core idea is to slide these filters across the sequence, detecting local patterns that are important for the task at hand. This allows the model to capture temporal dependencies and extract meaningful representations from the data. At its essence, the CNN 1D consists of convolutional layers, followed by activation functions, pooling layers for down sampling, and fully connected layers for object classification or regression tasks. This architecture has been efficiently successful employed in various applications, including speech recognition, natural language processing, and time series forecasting.

VI. RESULTS

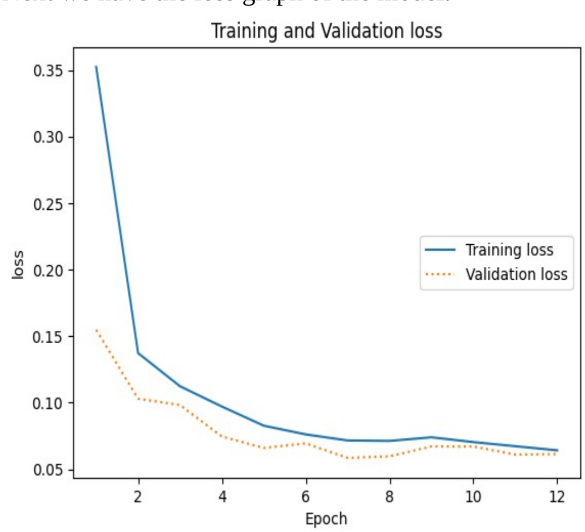
For CNN 1D:
Below is the table showcasing the accuracies captured by the model:

Accuracy	Recall	Precision	F1-Score
97.2970	97.2970	97.3335	97.2959

Next, we present the confusion matrix for the CNN 1D model's performance on the test set:



Next we have the loss graph of the model:



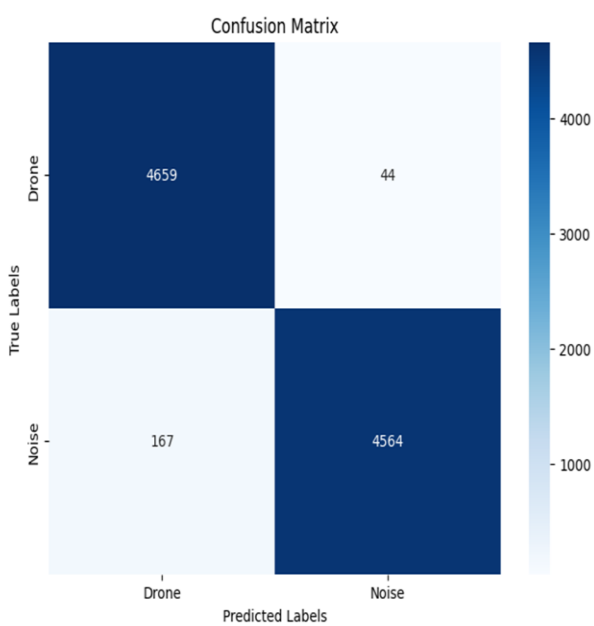
For CNN 1D:
The model we have used for the research purpose with the MFCC = 40 (MEL FREQUENCY CEPSTRAL COEFFICIENTS) is shown below

Audios	Feature	Duration	Accuracy	Precision	Recall	F-1 Score
Drone	MFCC=20	1 s	97.76	97.79	97.79	97.76
Noise						

So the model has 15 layers with 1 dense layer and 1 layer of 20% of dropout. And the result for the binary classification using Drone sound at different Sound to Noise ratio levels and Noise includes Highway, helicopter, bees, crowd, and rain.

The feature used is MFCC with number of coefficients equal to 20.

The confusion matrix for the same result recorded is shown below. There were several results recorded with different number of coefficients of MFCC the best result were given by MFCC = 20. The frame length equals to 1024 and hop length equals to 512.



Training and validation loss graph:

It confirms that the model is not overfitting

VII. DRONE CLASSIFICATION BASED ON ML

Research by Nijim and Mantrawadi [7] of drone detection through sound analysis. Their approach relied on the application of a Hidden Markov Model to identify DJI Phantom 3 and FPV 250 drones based on their sound profiles.

Here Jeon introduced a methodology encompassing Gaussian Mixture Model (GMM), Convolutional Neural Network (CNN), and Recurrent Neural Network (RNN) classification for drone presence detection within a 150-meter range [8]. To address the scarcity of acoustic drone data, the researchers proposed an inventive strategy involving the augmentation of assorted environmental sounds with drone audio to construct datasets. An intriguing facet of their study involved training and testing classifiers using distinct drone types. Their findings revealed that the RNN classifier yielded the highest performance (80%), followed by the GMM classifier (68%) and the CNN classifier (58%). It's noteworthy that the classifiers' effectiveness notably diminished when presented with previously unseen data.

Bernardini [9] employed a multi-class Support Vector Machine (SVM) classifier to distinguish drone sounds from other signals, including ambient crowd and daytime nature sounds. Their method encompassed the compilation of web audio data through an audio file scraper, emphasizing files with sampling rates exceeding 48 kHz. The curated dataset encompassed five 70-minute segments capturing sounds from flying drones, daytime nature, urban streets with traffic, passing trains, and crowds. These recordings were segmented into 5-second intervals for mediumterm analysis, with 20-millisecond sub-frames for short-term analysis, all featuring overlapping segments of 10 milliseconds. The authors extracted features such as short-time energy, temporal centroid, Zero Crossing Rate (ZCR), spectral centroid, spectral roll-off, and Mel Frequency Cepstral Coefficients (MFCCs) from the pre-processed signals, subsequently training an SVM classifier. The classification results demonstrated an accuracy of 96.4% in effectively detecting drone sounds amidst the other sound classes.

Seo [10] proposed an innovative approach involving the application of normalized Short-Time Fourier Transform (STFT) to generate 2D representations from acoustic signals emitted by drones. The sound signals were initially partitioned into segments of 20 milliseconds, incorporating a 50% overlap. Subsequently, the normalized STFT was extracted from these segments, serving as the input for a custom-designed Convolutional Neural Network (CNN).

The dataset employed in this study was derived from outdoor experimental measurements, capturing the acoustic profiles of hovering DJI Phantom 3 and Phantom 4 drones. This dataset encompassed a total of 68,931 sound frames from the drones, alongside 41,958 frames from non-drone sources. To assess the model's performance, evaluation took place on this dataset, following the introduction of Additive White Gaussian Noise (AWGN) to simulate real-world conditions. The most favorable outcomes emerged when training the CNN network over 100 epochs, specifically under low Signal-to-Noise Ratio (SNR) conditions. Under these circumstances, the model achieved a remarkable detection rate (DR) of 98.97% while maintaining a notably low false alarm rate (FAR) of 1.28%. This innovative methodology underscores the potential of normalized STFT in conjunction with CNNs for highly accurate drone detection in acoustically diverse settings.

VIII. CONCLUSION

In this research paper, exploration of highly accurate and precise models that can be used and applied for the classification purposes is successful. Many different combinations of techniques and their parameters explored to find the best and use the best.

Using the real-world scenario dataset and high-quality data augmentation the research contributes towards practical implementation more than the theoretical purposes. Best techniques and use cases from Machine learning and Deep learning are used.

Optimization of model parameters was the main-focus hence many different combinations of values are used so that overfitting and other degrading issues can be avoided. The implications of this research extend beyond the specific classification task at hand. The successful application of deep learning techniques in distinguishing drone sounds from background noise holds promise for broader use cases in the emerging field of acoustic surveillance and environmental monitoring. The insights and information gained from this study and research can contribute to urban planning and public safety measures, particularly in areas where drones may present privacy and security concerns.

Like any research, this research has some limitations. The effectiveness of the model may be influenced by the diversity of drone sounds encountered in different real-world scenarios. Further research could focus on increasing the diversity of the dataset to enhance the model's ability to generalize to varied drone sounds so that the model becomes more practically reliable and can be taken directly to use.

REFERENCES

[1] "Drones used to drop explosives on Jammu air base," Kamaljit Kaur Sandhu, Jammu, 2021.

[2] S. S. a. C. A. a. R. M. H. a. B. M. R. a. A. M. M. a. U. I. B. K. Y. a. J. Y. M. Alam, "RF-Enabled Deep-Learning-Assisted Drone Detection and Identification: An End-to-End Approach," *Sensors*, 2023.

[3] Z. S. a. M. H. a. C. Z. a. L. H. a. X. D. a. Y. Zhao, "Detection of LSSUAV using hash fingerprint based SVDD," *2017 IEEE International Conference on Communications (ICC)*, 2017.

[4] R. R. Fatemeh Mahdavi, Drone Detection Using Convolutional Neural Networks, ArXiv, 2021.

[5] M. S. a. S. D. K. a. N. S. a. M. Blumenstein, "A study on detecting drones using deep

convolutional neural networks," *2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS),* 2017.

[6] S. A. A.-E. a. A. K. A.-A. a. A. A.-A. a. A. Mohamed, "Audio Based Drone Detection and Identification using Deep Learning," 2019.

[7] M. a. M. N. Nijim, "Drone classification and identification system by phenome analysis using data mining techniques," in *2016 IEEE Symposium on Technologies for Homeland Security (HST)*, 2016.

[8] J.-W. S. Y.-J. L. W.-H. K. Y. K. a. H.-Y. Y. S. Jeon, "Empirical study of drone sound detection in real-life environment with deep neural networks," *CoRR*, 2017.

[9] A. a. M. F. a. P. E. a. C. L. Bernardini, "Drone detection by acoustic signature identification," *Electronic Imaging,* pp. 60-64, 2017.

[10] Y. a. J. B. a. I. S. Seo, "Drone Detection Using Convolutional Neural Networks with Acoustic STFT Features," in *2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, 2018, pp. 1-6.